

Controller Strategies for Automation Tool Use under Varying Levels of Trajectory Prediction Uncertainty

Susan Morey¹

San Jose State University/NASA Ames Research Center, Moffett Field, CA, 94035, USA

Dr. Thomas Prevôt, Ph.D.²

NASA Ames Research Center, Moffett Field, CA, 94035, USA

and

Joey Mercer³, Dr. Lynne Martin⁴, Nancy Bienert⁵, Christopher Cabrall³, Sarah Hunt⁵, Jeffrey Homola⁴, and Joshua Kraut³

San Jose State University/NASA Ames Research Center, Moffett Field, CA, 94035, USA

A human-in-the-loop simulation was conducted to examine the effects of varying levels of trajectory prediction uncertainty on air traffic controller workload and performance, as well as how strategies and the use of decision support tools change in response. This paper focuses on the strategies employed by two controllers from separate teams who worked in parallel but independently under identical conditions (airspace, arrival traffic, tools) with the goal of ensuring schedule conformance and safe separation for a dense arrival flow in en route airspace. Despite differences in strategy and methods, both controllers achieved high levels of schedule conformance and safe separation. Overall, results show that trajectory uncertainties introduced by wind and aircraft performance prediction errors do not affect the controllers' ability to manage traffic. Controller strategies were fairly robust to changes in error, though strategies were affected by the amount of delay to absorb (scheduled time of arrival minus estimated time of arrival). Using the results and observations, this paper proposes an ability to dynamically customize the display of information including delay time based on observed error to better accommodate different strategies and objectives.

I. Introduction

AS part of the overarching set of changes to the U.S. National Airspace System known as the Next Generation Air Transportation System (NextGen), the Federal Aviation Administration (FAA) aims to develop and implement automation aids to support human air traffic controllers in meeting the forecasted growth in air traffic demand. Such automation aids depend upon aircraft trajectory predictions made by air traffic control ground automation systems [1]. Trajectory predictions inherently have some uncertainty. Being trajectory-based, the automation tools used in this simulation were subject to various intentional errors, meaning the controllers were interacting with less-than-perfect tools. This approach made it possible to examine the point at which the automation tools would become unacceptable to the controllers, and the system performance was no longer meeting separation and metering objectives. The varying levels of uncertainty were intended to mimic the deployment of imperfect automation systems. If it were shown that there is an ideal band of accuracy, below which the tools are not useful,

¹Research Engineer, San Jose State University Research Foundation, NASA Ames Research Center, AIAA Member.

²Research General Engineer, NASA Ames Research Center, Moffett Field, CA, AIAA Senior Member.

³Research Psychologist, San Jose State University Research Foundation, NASA Ames Research Center, Moffett Field, CA, AIAA Member.

⁴Research Psychologist, San Jose State University Research Foundation, NASA Ames Research Center, Moffett Field, CA.

⁵Research Assistant, San Jose State University Research Foundation, NASA Ames Research Center, Moffett Field, CA.

and above which the tools are of minimal additional benefit, then the FAA and industry could develop automation systems to the required accuracy level.

Research in human-automation systems has shown that when it comes to using automation to support human tasks in a dynamic environment such as air traffic control, strategy and individual difference matter. Kirlik [2] investigated an autopilot tool in which participants employed the tool to support pilot goals, but no participant did so in the manner intended by the designers. Further examination revealed a participant-specific tool-use strategy based on a combination of environmental and design constraints. Kirlik concluded that system performance alone is not a guarantee of human-automation system success. In addition to an accurate system, designers must understand and account for the potential environmental changes and influences that will affect operator strategy. This assessment was taken one step further by Casner [3] with real world pilot interfaces and routes. In an effort to understand the range of potential context effects, Casner quantified aspects of a pilot's complex environment and looked for correlations in tool use as applied to problem solving. Results suggested that i) strategies reflect environmental demands which are tuned to a specific context, especially if there are predictable features in the environment, and ii) even when a tool (in this case the Flight Management System) is preferred, there are environmental scenarios where its use is unfeasible and it must be set aside. Casner's study is consistent with Kirlik's conclusion that the decision to set aside a tool is as important as that to use it for optimal performance.

Strategy and individual differences must be considered when developing controller automation tools, so that the tools help the controller achieve his or her goals. This paper presents data on controller strategies for using trajectory-based automation tools under uncertainty. These data were gathered during a human-in-the-loop simulation that examined the ability of air traffic controllers to deliver aircraft on time to a meter fix and maintain separation in varying levels of automation system accuracy. Two approaches to using automation tools to achieve schedule conformance at a meter fix were analyzed. Strategies for how the controllers used the available automation tools and implications for tool design are discussed.

This experiment investigated whether increased tool accuracy results in improved airspace management, efficiency, or safety. Participant controllers were presented with an en route arrival problem in which they sought to deliver arrival aircraft to a specific point (the meter fix) at a particular time, while providing standard separation services. Arrivals were scheduled using a trajectory-based management scheduler, and the controllers were asked to deliver aircraft in sequence and according to the schedule. Trajectory prediction uncertainty was introduced into the simulation, and the effects of the uncertainty on metering precision was measured. While the cause of the induced error (wind forecast and aircraft performance) in this simulation was not directly attributable to the tools, the error was expressed to the controller through their tools' varying levels of accuracy.

A. Focus on Strategies and Tool Design

To examine the effects of trajectory prediction uncertainty on sector performance and automation usage, a simulation was conducted over the course of one week in which controller teams participated in two simulations conducted simultaneously and in parallel. These two independent instances of the experiment (i.e., two "worlds") enabled a method for comparing how two controllers independently tasked with the same airspace, goals and tools naturally used noticeably different strategies, as evidenced by the types of clearances issued to achieve schedule conformances. The two strategies varied in easily observed ways despite identical traffic scenarios, experiment conditions, and objectives.

This paper examines A) whether observed differences in approach between the two worlds were representative of consistently applied strategies, B) how successful the two strategies were in meeting the goals of schedule conformance and safety, C) how the strategies were affected by the experiment conditions, and D) how these strategy insights can lead to improved automation tools. Since the participants were free to use any strategy they wanted (and were not asked explicitly to focus on strategy as part of the study), choice of strategy cannot be separated here from individual differences. It is intended that research findings from this analysis will support automation enhancements to operational air traffic control systems.

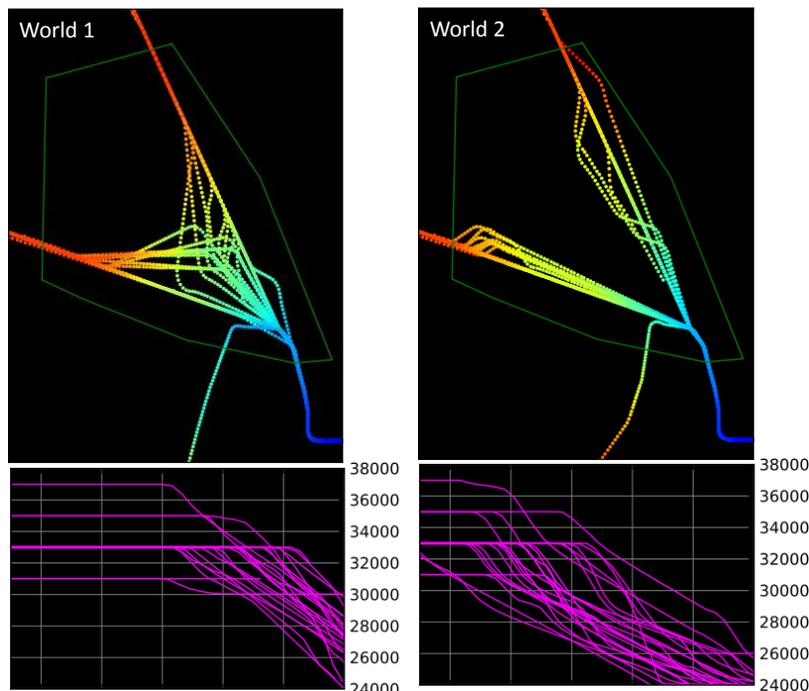


Figure 1. Plots of Aircraft Lateral and Vertical Paths Across Worlds.

Figure 1 depicts the lateral and vertical paths flown by aircraft in one of the experiment runs, and illustrates the differences in controller strategy observed during the experiment between the controller in World 1 (Controller 1) and the controller in World 2 (Controller 2). Aircraft flight paths from each of the two independent simulations are compared side-by-side, with World 1 on the left and World 2 on the right. Approximately 25 arrival flights traversed the test airspace during each run. For each world, the upper plot shows the lateral paths of the arrival traffic in the run, where the color change indicates altitude. The lower plot shows the vertical paths of the aircraft with altitudes shown on the y axis, and the x axis represents distance to the meter fix. Controller 1's use of path stretching by route amendments is visible by comparing the two images. Likewise, Controller 2 descended aircraft earlier, as evidenced by the color change in the upper plot, and by the earlier descents in the lower plot.

2 descended aircraft earlier, as evidenced by the color change in the upper plot, and by the earlier descents in the lower plot.

The rationale for the analysis described here is to understand more about these seemingly very different approaches to metering arrival traffic, including how these approaches influenced controllers' use of automation tools.

II. Experiment Methods

A. Participants

All test participant controller positions were staffed by retired radar-certified Oakland Air Route Traffic Control Center (ARTCC) controllers. The two controllers who staffed the sectors analyzed here were both retired in 2007 after 25 years of experience. Additional retired radar-certified controllers managed the airspace around the test sectors as confederates. Licensed pilots, most of whom were university aviation students, served as confederate pseudo-pilots, controlling multiple aircraft in the simulation environment.

B. Environment

The airspace represented included ZTL05 (low altitude) and ZTL06 (high altitude) en route sectors over Atlanta, shown in Figure 2, with one experimental controller position for each sector. The arrivals into Atlanta flew through ZTL06's airspace, and then continued their descent into ZTL05's airspace and across the meter fix ERLIN, to which the aircraft were being scheduled. Atlanta arrivals flew from the northwest through NEUTO, FIXIK, etc. or from the west through CALCO, BAKRD, etc. and then descended via the RPTOR1 arrival. A small number of aircraft came from the southwest through SLAKR to join the arrival flow. Additionally, the sectors had overflight traffic for realism. The experiment introduced trajectory prediction uncertainty (error) in two forms and analyzed the interaction between those errors and controller performance.

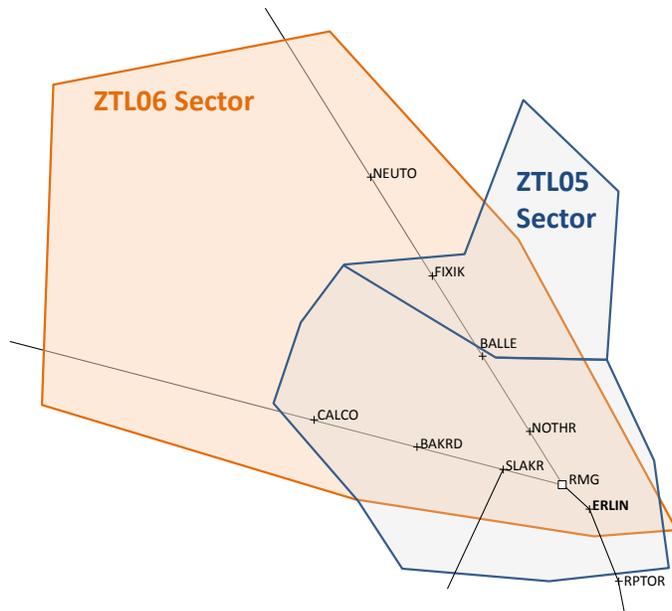


Figure 2. Experimental Sectors.

concepts that would be used during the experiment runs. Each experiment run lasted approximately 55 minutes. After each run and condition, the controllers completed a questionnaire. After the last experimental run, controllers completed a more exhaustive post-simulation questionnaire, which included questions related to strategy and tool use.

Participant controllers were asked to deliver arrival aircraft according to a fixed schedule at the meter fix, while providing standard separation services. Each arrival aircraft was assigned a Scheduled Time of Arrival (STA) over the meter fix ERLIN by the scheduler, setting this fixed time and aircraft sequence based on the current Estimated Time of Arrival (ETA) at ERLIN of the aircraft and other aircraft. Additionally, and with varying levels of uncertainty, the automation tools continuously re-calculated an Estimated Time of Arrival (ETA) over ERLIN. At any given point, subtracting the ETA from the STA yields a measure of schedule conformance, called delay, for an aircraft, that indicated how early or late that aircraft was at that point in time relative to the schedule. Positive delay times indicate that an aircraft is estimated to cross ERLIN early and therefore needs to have delay absorbed in order to arrive on time. A negative delay time is associated with an aircraft that is estimated to arrive late at ERLIN, unless the controller issues a speed increase to the pilot. Delay precision was shown to the tens of seconds, and the controllers were asked to deliver the aircraft in sequence over the meter fix within a +/- 25 second delay window.

In each of the two worlds, a single experimental ZTL06 controller participated. The over-flight aircraft in the ZTL06 sector originated from a variety of directions, thereby increasing the complexity and realism of the sector operations. The arrival aircraft originated from the Northwest and West along the routes displayed in Figure 2. These aircraft, according to the discretion of the controller, began their descent from cruise altitude inside the ZTL06 sector, before being handed off to the ZTL05 sector. All clearances in the experimental runs were issued by voice.

D. Experiment Design

In this experiment the independent variables were the two sources of trajectory prediction uncertainty: the difference between forecasted and actual winds, described as wind error, and the difference between modeled and actual aircraft performance, described as aircraft performance error. The source and level of these uncertainties were outlined to participants at the beginning of each simulation run, and the runs were mostly ordered such that they gradually increased in uncertainty. The experiment was designed as a partial matrix with three aircraft performance error levels and five wind forecast error levels, resulting in seven combined experiment conditions.

The aircraft performance errors were introduced as a scaling factor in the simulation software. A range of aircraft performance errors were distributed across all aircraft in the traffic scenarios such that two standard deviations (95%) of the aircraft had errors within +/- 5% in the realistic error condition, within +/- 25% in the large error condition, and a condition with no error. These errors affected the way the simulated aircraft flew, such that they descended at a rate faster or slower than the trajectory model predicted.

A simulation world was created using the Multi Aircraft Control System (MACS) software to simulate ground-side and air-side operations in the U.S. National Airspace System [4]. En route equipment and systems used by the controllers during the experiment were chosen to closely replicate the look and feel of equipment used at FAA en route air traffic control facilities. The traffic scenarios were modified from historical real-world traffic, wind, and route data to meet requirements of the experimental design. Voice communication equipment was utilized, enabling voice communication between controllers and pseudo-pilots.

C. Procedure and Task Description

Participant controllers were given two days of training during the week prior to the experiment. They gained familiarity with their sector's traffic patterns, rules, and requirements, as well as with the tools and

Wind forecast error was introduced by creating differences between the forecast and actual winds, and varied from 0 knots (i.e., no error), 10 knots, 20 knots, 30 knots, and 40 knots. The wind and aircraft performance errors were used by the automation to build trajectory predictions to drive automation aids. Since the underlying trajectories were subject to the prediction errors, the tools displayed inaccurate information to the controllers. Delay information, for instance, requires a trajectory prediction. If the trajectory prediction were calculated using inaccurate information, the ETA for an aircraft would also be inaccurate, and may mislead the controller into applying control actions that do not improve the position of the aircraft relative to its STA. In theory, larger error conditions would result in larger discrepancies between an aircraft's ETA and its actual arrival time.

There were three traffic scenarios used in the experiment, referred to as A, B, and C. Traffic scenarios A and B were very similar in terms of traffic count, complexity, and delay times needed. In traffic scenario C, the arrival aircraft required higher amounts of delay; runs using traffic scenario C are accordingly excluded from this analysis.

Table 1. Experiment Matrix Showing Runs Analyzed.

Run	Experiment Condition (aircraft performance error, wind forecast error)	Traffic Scenario	Wind Error Bias (Forecast – Environment)
1	NN (none, 0 knots)	A	None
2	NN (none, 0 knots)	B	None
3	RR (+/-5%, 10 knots)	A	10 (80-70)
4	RR (+/-5%, 10 knots)	B	-10 (80-90)
5	RL (+/-5%, 30 knots)	A	30 (100-70)
6	RL (+/-5%, 30 knots)	B	-30 (60-90)
7	LR (+/-25%, 10 knots)	A	10 (80-70)
8	LR (+/-25%, 10 knots)	B	-10 (80-90)
9	RM (+/-5%, 20 knots)	A	-20 (60-80)
10	RM (+/-5%, 20 knots)	B	20 (100-80)
11	LL (+/-25%, 30 knots)	A	30 (100-70)
12	LL (+/-25%, 30 knots)	B	-30 (60-90)
15*	LXL (+/-25%, 40 knots)	A	40 (100-60)
17*	LXL (+/-25%, 40 knots)	B	-40 (60-100)

*Runs 13, 14, 16 excluded from this paper's analysis; run 13 was a repeat of run 11, and runs 14 and 16 used a different traffic scenario.

need in order to reach the destination at the STA, appearing to controllers and the automation that the aircraft had a large delay to be absorbed. Under-predicted wind forecast errors meant that for an aircraft flying with a 70 knot tailwind, the automation used an estimate of a 60 knot tailwind. This resulted in a later ETA, and thus it appeared to controllers that the aircraft needed less delay than it would actually need to arrive at its destination on-time.

Data were collected to quantify how controller strategies differed by world, scenario, experiment condition, and positive/negative wind bias error. Data collected included clearances issued, tool-related events, aircraft states and schedule states at certain points in time, for arrival aircraft that are worked by the test sectors and crossed the meter fix. Additional data were collected to examine the success of the controllers and their strategies, including data on schedule conformance, flight path length, and losses of separation.

E. Controller Tools

The controllers were provided tools for separation assurance and schedule conformance. Some of these tools emulated today's fielded automation tools, and some were research prototypes implemented only in the MACS simulation environment. Separation assurance tools available to controllers included a conflict list similar to today's User Request Evaluation (URET) conflict list, and a conflict probe which showed the time to estimated loss of separation in the data block. The schedule conformance tools available to controllers included a delay time by the aircraft symbol, a meter list with delay times for each aircraft in sequence, and a series of trial plan functions that allowed the controller to query the automation for a solution to the delay or advance required in a certain dimension (e.g., speed, altitude or route). Figure 3 shows the controller tools available to support the separation assurance and schedule conformance tasks.

Seventeen data collection runs were conducted over four days, with four baseline data collection runs conducted in the prior week. Table 1 lists the experiment runs that were included in this paper's analysis. Mercer [5] provides an overview of results regarding the primary experiment conditions for all runs.

A post-hoc condition was effectively introduced by the traffic scenarios used. Except for the Moderate condition (RM in column 2 of Table 1), traffic scenario A had over-predicted wind forecast errors (thus showing higher delay times to controllers) and the traffic scenario B had under-predicted wind forecast errors (thus showing lower delay times to controllers). Over-predicted wind forecast errors meant that for an aircraft flying with a 70 knot tailwind, the automation predicted it had an 80 knot tailwind. This positive wind error bias caused the automation to display a larger delay than the aircraft would actually

Table 2. Trial Planning Tools.

Tool	Tool Purpose	Behavior & Use
TT: Manual Trial Planner	Used to bring up trial plan graphics (a trial plan route line that can be manually dragged to display route amendment).	Shows conflict and expected delay time based on manually dragging route. Used for separation and metering.
TA: Altitude Trial Plan	Brings up trial plan graphics with a climb or descent to the entered altitude. Controller enters TA <altitude, e.g. 240> <aircraft>.	Shows conflict and delay time impacts based on entered altitude. Used for separation and metering.
TR: Route Trial Plan	Brings up trial plan graphics based on the entered waypoint. Controller enters TR <fix name, e.g. BALLE> <aircraft>.	Shows conflict and delay time impacts based on entered route. Used for separation and metering.
TS: Speed Trial Plan	Brings up trial plan graphics with a proposed speed, if controller enters TS <aircraft>. Controller may also use the trial planning tools to test the effects of a specific speed, such as TS 250 <aircraft>.	Provides a resolution for metering; proposes a conflict-free speed to get expected delay as close as possible to zero, where aircraft ETA matches STA, within parameters based on aircraft performance and current altitude. Used for metering.
TM: Metering Trial Plan	Brings up trial plan graphics with a proposed speed, route, and/or altitude. Controller enters TM <aircraft>.	Provides a resolution for metering; proposes a conflict-free resolution to get delay as close as possible to zero, within parameters based on aircraft performance and current altitude. Used for separation and metering.

In this experiment, the controllers primarily used the trial planning tools to issue clearances to achieve schedule conformance, though they were also occasionally used for separation assurance. When a clearance was issued and amended into the system, the delay time immediately reflected the impact of the clearance, as opposed to a gradual change in the delay value as the aircraft continued along its path. Although encouraged, use of the tools was at the controller's discretion, providing an opportunity to examine when and how these tools were used in the various experiment conditions. While the rest of the tools were optional, the controllers effectively had to use the meter list in order to know the sequence of the aircraft over the meter fix.

III. Results

Overall, trajectory prediction uncertainty driven by aircraft performance error and wind forecast error was not found to affect controller performance as measured by schedule conformance at the meter fix and separation losses [5]. Although originally intended to be simple variants, the two traffic scenarios A and B resulted in aircraft entering the test sector ZTL06 with more delay to absorb in traffic scenario A than in traffic scenario B. Accordingly, the results are analyzed separately by traffic scenario. Primary results from the Trajectory Prediction Uncertainty simulation are reported in Mercer [5]. An additional analysis of strategy and tool use during high demand situations is reported in Kraut [6].

Data presented in this section include statistically significant results where indicated, and observed results in cases where data were insufficient to test significance. Only aircraft that traversed the test sectors and crossed the meter fix by the end of the simulation run were included in the analyzed data. Given the applied nature of this experiment, some discussion of results and their importance in terms of strategy and tool use is included in this section. It is important to understand that co-varying elements are present in the data. Comparative analyses will serve only as indications to possible outcomes: outcomes that will require controlled testing to validate.

A. Controller Strategies

Differences in controller strategies for arrival metering were apparent in watching the controllers independently perform the same metering task. The high altitude sector ZTL06 was chosen here for an analysis of metering strategy and automation tool usage for a variety of reasons. First, most of the delay was absorbed in ZTL06 rather than in ZTL05. Second, during the experiment, researchers observed very clear differences in strategy between the

two ZTL06 controllers. ZTL06 controllers reported higher and more variable workload ratings than the ZTL05 low altitude sector controllers did. Finally, since aircraft first pass through ZTL06, the ZTL05 controller's performance was influenced by the performance of the ZTL06 controller, so it was not possible to assess ZTL05 sector performance independently of ZTL06 sector performance.

1. Delay Absorption

Figure 4 shows delay times at the entrance to sector ZTL06 (when ZTL06 accepts the handoff) and when the aircraft is handed off from ZTL06 to ZTL05 (ZTL05 accepts the handoff). These delay values were the values shown to the controller and were affected by the introduced error. This figure shows a clear difference between traffic scenario A and B in the amount of delay displayed to the controller at the time that ZTL06 accepts the handoff ($F = 1132, df = 1, 12, p < .001$).

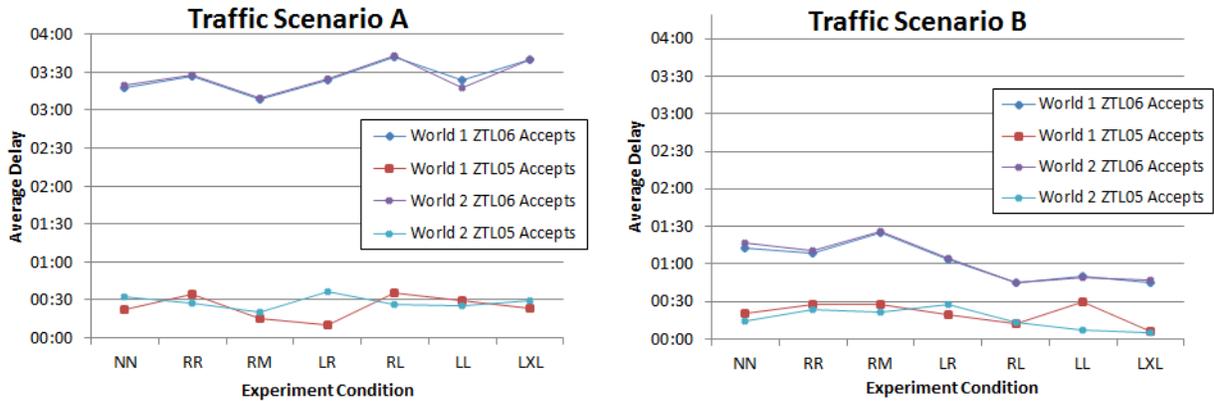


Figure 4. Average Delay Times Entering ZTL06, Exiting ZTL06 and at Meter Fix by Traffic Scenario, Experiment Condition, and World.

In each experimental condition and in both worlds, however, the average delay at handoff to ZTL05 was comparable and consistent, and always less than 40 seconds, indicating that in both worlds and in both traffic scenarios, controllers absorbed the majority of the delay and handed off the aircraft to ZTL05 with about the same remaining delay. No obvious trend was seen with regard to the delay absorbed by experiment condition.

Table 3. Clearance Count by World and Traffic.

World	Total Clearances	Traffic Scenario A	Traffic Scenario B
1	1055	633	422
2	1386	842	544

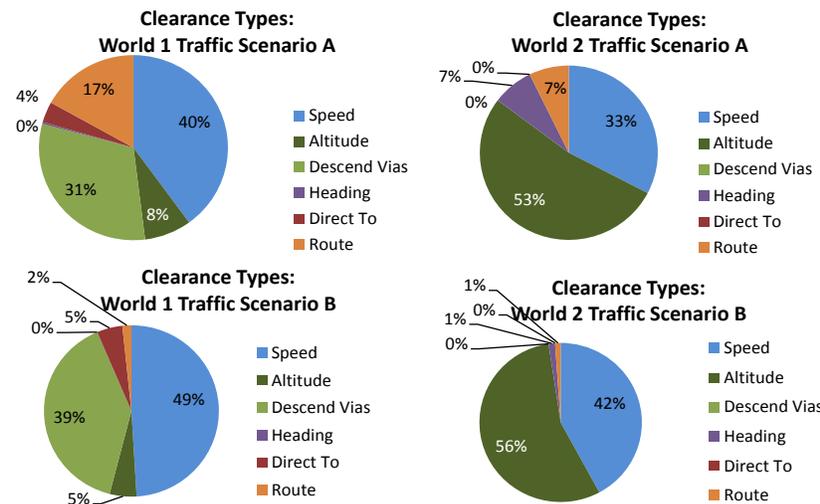


Figure 5. Clearance Types Issued by World and Traffic.

2. Clearance Analysis

Analysis of the clearances issued by ZTL06 controller, which were issued primarily to absorb delay, indicated that Controller 2 issued 30% more clearances than Controller 1 with both traffic scenarios, as shown in Table 3 ($t = -2.995, df = 26, p = .006$). When looking at the data by traffic scenario, having to absorb greater delays in traffic scenario A caused both controllers to issue more clearances, about 50% more for both controllers ($t = 6.281, df = 13, p = .001$).

Types of clearances given were also analyzed. Clearances for delay absorption were counted by type: speed (e.g., "maintain 250 knots"); altitude (e.g., "descend and maintain flight level 240"); pilot's

discretion / descend via clearances (e.g., “descend via the RAPTOR1 arrival”); heading (e.g., “turn left heading 065, vector for spacing”); direct to (e.g., “clear direct RMG and the remainder of the RAPTOR1 arrival”); and route (e.g., “clear direct BALLE direct RMG”). Figure 5 shows that the controllers issued different clearance types other than speed and route ($t = -12.114$ (altitude), 39.216 (pilot’s discretion altitude), -4.147 (heading), 8.653 (direct to), $df = 26$, $p < .001$ in all cases).

Types of clearances issued by experiment condition were also investigated. Figure 6 displays the total count of clearances by clearance type and by experiment condition. The plots are arranged by traffic scenario from top to bottom, and by World (i.e., controller) from left to right. The experiment conditions are on the x axis, ordered from least error to most error. When separated by condition, due to having only two runs per condition, data were insufficient to quantitatively determine differences. Figure 6 appears to show that as errors increase right to left, clearance count increases for traffic scenario A but not for traffic scenario B; in runs with traffic scenario A, the

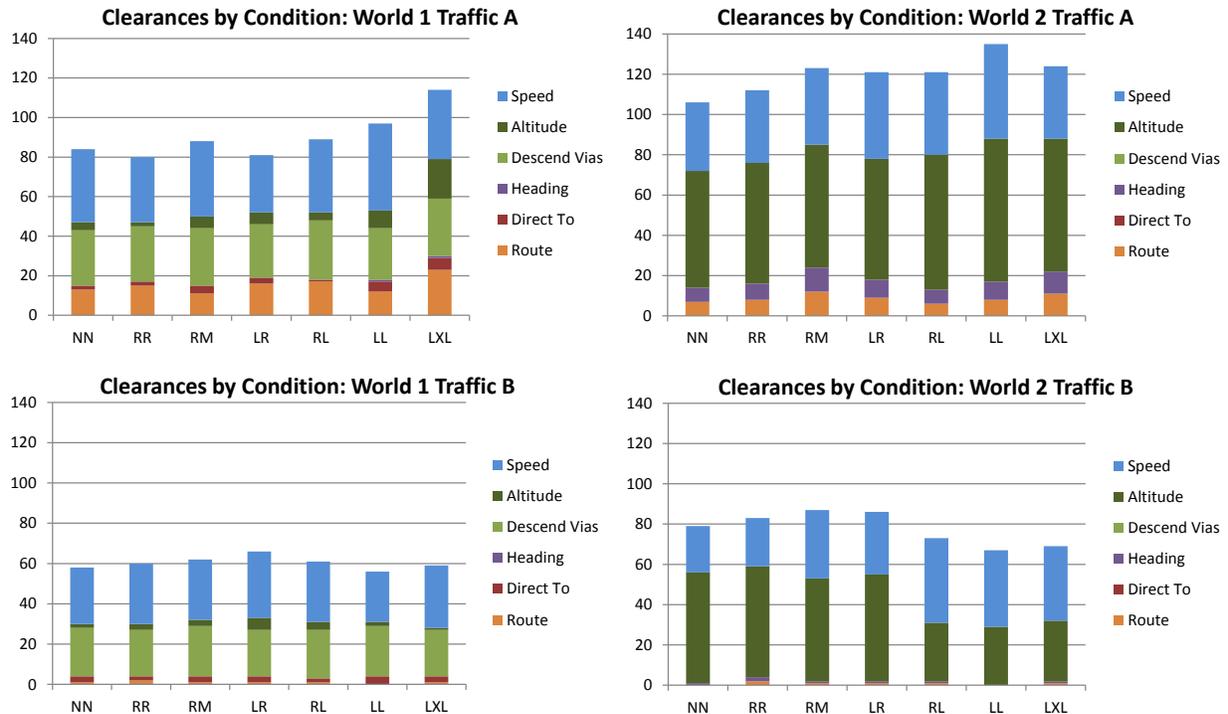


Figure 6. Number of Clearance Types Issued by Experiment Condition, World and Traffic Scenario.

automation displayed larger delay values.

3. Tool Usage

In the questionnaire responses, controllers reported using primarily the same tools, namely the conflict list, conflict probe, delay time in the datablock, meter list, manual trial plan, and speed trial plan tool. They both reported using all those tools frequently across all experiment conditions. Controller 2 also reported using the altitude trial plan tool for altitude resolutions. The three tools the controllers reported using the most were the speed trial plan tool, the meter list, and the delay time in the datablock. Responses show that the controllers rated the tools they used the most as most accurate and most stable. When asked which tool they would choose if they could select just one, both controllers selected the speed trial plan tool.

The controllers were asked in post-condition questionnaires how often they directly used the information provided by the tools, and how often they used the information indirectly. Using the tool’s information indirectly refers to factoring tool information into their decision making, but not using the information as-is, such as considering proposed speeds suggested by the automation, but issuing something slightly different. Both controllers reported that they used tool information both directly and indirectly in all conditions; the error never increased to the point that they reported giving up on using any of the tool information directly. Overall, they reported using the information indirectly more often than directly (i.e., nine times controllers reported using the information indirectly at least ‘three-quarters of the time’ and five times controllers reported using the information directly at least ‘three-quarters of the time,’ with no response given by Controller 2 for one experiment condition, RM). When comparing direct and indirect use of tool information by experiment condition, a slight pattern emerged for Controller 2: as the

error level increased from NN to LL, the controller’s direct usage rate decreased, suggesting they used the tool information less directly as its accuracy decreased. Controller 1 showed no such pattern, indicating that his use of tool information (either directly or indirectly) was not affected by experiment condition. Data were insufficient to determine if direct and indirect tool information use varied by traffic scenario.

B. Measures of Success

1. Schedule Conformance

Schedule conformance was defined as the percentage of aircraft that crossed the meter fix laterally within 25

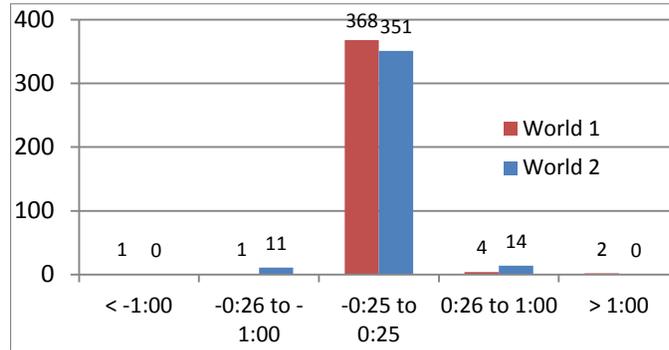


Figure 7. Delay Times at Meter Fix by World.

seconds of the STA and meeting the altitude restriction within 200 feet. Due to the location of the meter fix, schedule conformance reflected the combined efforts of both sector controllers in a world, that is, World 1 data reflected aircraft managed by both World 1 ZTL05 and World 1 ZTL06 controllers. Overall, both strategies were very successful in meeting schedule conformance at the meter fix. Figure 7 shows the number of aircraft in each world that crossed ERLIN with delay times within +/- 25 seconds, between 1 minute and 25 seconds, and exceeding +/- 1 minute. The data show that both controllers were quite successful: World 1 achieved a 98% overall success rate, and World 2 achieved a 93% overall success rate. However, while the data showed each world achieved a high schedule conformance success rate, World 1’s success rate was significantly higher ($\chi^2(1) = 12.080, p = .001$), as assessed by a Chi Square Test of Association where all cells had an expected count greater than five.

Table 4 shows the delay times at the meter fix, separated according to world and experiment condition. The data appear to indicate that experiment conditions did not hamper controllers’ ability to deliver aircraft on time over the

Table 4. Delay Times at Meter Fix by World and Condition.

Delay at ERLIN	World 1:World 2						
	NN	RR	RM	LR	RL	LL	LXL
more than 1 min late	0:0	0:0	0:0	0:0	1:0	0:0	0:0
1 min to 25 secs late	0:3	0:2	0:0	1:1	0:3	0:1	0:1
within +/- 25 secs (met goal)	54:48	54:49	54:54	52:51	51:46	50:51	53:52
1 min to 25 secs early	0:3	0:3	1:1	0:2	0:3	3:2	0:0
more than 1 min early	0:0	0:0	0:0	1:0	0:0	1:0	0:0

meter fix. Note that aircraft that did not cross the meter fix within geographic parameters are not included. In World 2, more aircraft arrived outside of the 25 second window, but none more than 1 minute away from their STA.

2. Flight Path Length

Distance flown was used as an indicator of the degree to which controllers efficiently managed the traffic. Measurements were taken between a 200 nmi arc around the meter fix and the meter fix itself. Under this analysis, an aircraft flying uninterrupted directly to the meter fix would have flown exactly 200 nmi. Analysis of flight path length across worlds did show differences, which neared statistical significance ($F(1, 358) = 3.528, p = .061$), with a mean of 210.6 nmi in World 1 and a mean of 208.3 nmi in World 2. Like the schedule performance data, flight path efficiency data were considered from the perspective of the collective efforts of both sectors.

3. Losses of Separation

The data were also analyzed for the occurrence of separation violations (both operational errors and proximity events). There was one loss of separation in the runs considered here, which occurred in World 1 during run 11, an LL experiment condition using the traffic scenario A (the higher delay time traffic), and involved an arrival aircraft and an over-flight aircraft. At its closest point, the planes were 594 feet apart vertically and 3.6 miles apart laterally.

IV. Discussion

A. Observations of Strategy

To identify the strategies employed by the Controller 1 and Controller 2, screen capture videos with verbal clearances were reviewed by the first author, including comparing controller handling of a subset of aircraft over each arrival flow and across two experiment conditions. Observations from the video analysis are consistent with the analytic data on clearances issued and questionnaire responses on tool usage. Arrival aircraft are only in the ZTL06 sector for about five minutes, which does not allow much time for controllers to absorb the three-plus minutes of delay commonly seen in traffic scenario A before handing the aircraft off to the ZTL05 controller.

From observing the experiment initially and after reviewing a subset of the runs, clear patterns for absorbing delay were seen in each world. Regularly, both controllers issued speed clearances early, and then used other methods as needed to absorb more delay. In World 1, the controller first issued a speed reduction and then issued a clearance for the plane to descend via the arrival route. A minute or two later, if there was still more delay he wished to absorb, he used the manual trial plan tool to determine and issue a route amendment.

In World 2, the controller issued clearances as soon as he took the handoff, in some cases 20 to 25 miles outside his sector boundary, and almost always started by issuing a speed clearance. He also issued immediate altitude clearances at the same time as the speed reduction or soon after, depending on the amount of delay to absorb, such as for delays of over three and half minutes. He used the speed trial plan tool to get a speed recommendation, and then would manually drag the routes in some cases to determine a route (particularly early in the run). He would amend the system with the route but verbally issue a heading to the aircraft. He would leave the aircraft on the heading until the desired amount of delay was absorbed, then issue a route amendment to a fix on the arrival route. Amending the system with the route caused the delay time to update based on the route. In other cases, Controller 2 issued speed and then reduced altitude so that the speed decrease was more effective in absorbing delay, but did not issue any headings or route amendments. Controller 2 seemed to use the trial plan tools to generate ideas, amend the system with the desired trial plan route, but then issued heading changes to the aircraft.

Both controllers used the trial planning tools more in the early parts of each experimental run; for later aircraft, they would issue the speed clearances without trial planning, based on what they had used earlier. The controllers reported working with their ZTL05 sector controllers to set a target delay time at handoff that was often higher than +25 seconds delay to compensate for the error, leaving some buffer for the ZTL05 sector controller to work with. Controllers in World 2 reported collaboration between ZTL06 and ZTL05 in each run while the controllers in World 1 reported collaborating in 50-75% of the runs. Controllers used the speeds recommended by the automation more or less directly depending on the wind forecast error. To account for the error, they adjusted their target delay time range at handoff to the next sector (e.g., +30 to +50 seconds, or -10 to -20 seconds) based on their understanding of the wind error. The controllers often used different target delay time ranges for the west and northwest arrival routes to account for wind direction.

The finding that there were frequently delay amount agreements made between ZTL05 and ZTL06 controllers in the different worlds is an interesting result of this experiment. Unfortunately, these agreements between ZTL06 and ZTL05 controllers were not consistently documented during the experiment, making in-depth analyses difficult. An analysis of the available data suggests that the two worlds developed different styles of collaboration. In a questionnaire response, the World 1 ZTL06 controller stated that “Once the problem starts and the first few airplanes go in [to the meter fix ERLIN], the low sector lets me know what kind of a time pad is needed for each route (+10 to +30 or 00 to -10, etc.)” From observing each pair, it seemed that the World 2 controllers were collaborating on the desired delay amount at the handoff to ZTL05, and would often collaboratively determine delay amounts for each arrival route. In contrast, the World 1 controllers discussed the desired delay on an as-needed basis, when the ZTL05 controller wanted the ZTL06 controller to change his approach. Follow-on studies should explicitly seek out information about the collaboration styles, approaches, and delay time agreements made between sectors.

B. Interpretation of Results

The amount of delay that the controllers perceived they had to absorb affected their strategies, which is why traffic scenario mattered. When controllers perceived that they had more delay to absorb, they issued more clearances and used different types of clearances. This is evidenced by the finding that in both worlds, controllers absorbed more delay before handoff to ZTL05 in runs with more perceived delay, occurring in runs with traffic scenario A. Data were analyzed to isolate whether it was the positive or negative wind error bias (forecast winds minus actual winds) that had an impact on controller strategy and performance or whether it was the scenarios themselves; the data showed that the traffic scenarios themselves differed.

The two controllers' strategies were different from one another, but both were very effective. Controller strategies differed in numbers of clearances issued, types of clearances issued, and in ways of using the tools (e.g. using route trial plan directly or indirectly). Controller 2 issued more clearances overall than Controller 1 did. Controller 1's strategy relied more on direct use of the tools, and he used the trial planning tools to issue route changes. He also used more pilot's discretion (i.e. "Descend Via") clearances. Controller 2 issued immediate altitude clearances and used vectoring / heading changes as needed to absorb additional delay. Often Controller 2's vectoring was influenced by route trial plan tool information, but he less commonly issued route amendments to the aircraft to absorb delay, preferring to issue vectors and then have the aircraft rejoin the route once sufficient delay was absorbed. These differences were reflected in the resulting flight paths, both laterally and vertically, seen visually in the plotted aircraft paths (Figure 1) and in the flight path length analysis. Controller 2's strategy lowered aircraft earlier to absorb delay more effectively with speed rather than issuing lateral route changes, resulting in shorter average path lengths. Both strategies resulted in very high schedule conformance at the meter fix.

Controller strategies did change based on the amount of delay they perceived they had to absorb, here driven by the traffic scenario used. When presented with more delay to absorb, both controllers issued more route amendments. Controller 1 also issued more altitude amendments with higher delays, and Controller 2 issued more headings when presented with more delay.

One of the main findings of the experiment is that even large tool inaccuracies, when consistent, can be tolerated by the controllers and will not necessarily render the tool ineffective or cause the controllers to abandon use of the tool. As described by the controllers, "once you understood the errors, you adapted and it was fine," which suggests that the high levels of error introduced in this simulation did not sufficiently challenge their performance. They described it as "figuring out the problem" and one controller reported he could figure out the problem after watching the first two aircraft traverse the meter fix. Controllers were able to adapt their use of their preferred tools to the different experiment conditions, with one of the two controllers using the tool information more indirectly as experiment conditions increased. Tools that can be used in a wider variety of ways will accommodate a broader range of controller strategies, and will remain usable across varying environmental conditions, such as delay time to be absorbed.

This experiment's results are consistent with Kirlik's [1] findings that environmental factors influence and modify controller strategy and that tools must be designed to be adaptable to differences in strategy, and as environmental factors change. An example of the broad range of ways a single tool is used by controllers is seen in this experiment with the route trial plan tool. Controller 2 used the route trial plan tool to find a route amendment and entered it into the system to update the delay time, but then issued a heading to the aircraft. Clearly, the route trial plan tool was giving him ideas of how much vectoring to do with the aircraft, and he entered it into the system so that the delay time would update, but he kept the aircraft on a heading so that he could work it more precisely and in a flexible manner, turning the aircraft back on route when the desired amount of delay had been absorbed. Controller 1 used the route trial plan tool directly, manually adjusting the trial plan route line as desired.

C. Implications for Tool Design

Results from this experiment emphasize the need for tools to be flexible enough to handle a wide range of strategies that vary based on individual differences, and which may change due to changes in environment. There were three tools reported to be used consistently by both controllers (meter list, delay time in the datablock and the speed trial plan tool). The latter two were optional (the meter list had the order of aircraft and STAs and therefore was required for any metering to take place), and were used by both controllers despite their having very different strategies. The instant update to the delay time shown in the data block and the automation-provided speed recommendations are strong candidates for inclusion in future en route automation enhancements, because their usefulness persisted across varying delay levels, aircraft performance errors, and wind errors. The speed trial plan tool can be used either to request a suggested speed from the automation, or to probe a controller-provided speed for conflict and impact on delay time: both methods were used by the controllers in this study. The controllers sometimes referred to the speed trial plan tool interchangeably with the behavior of the delay time in the datablock updating immediately due to an amendment. Therefore, the high rating of the trial plan speed tool should be interpreted as both a vote for the immediate feedback in the delay time and for the tool's speed recommendations to meet the schedule conformance goal at the meter fix.

The controllers set goals to have specific non-zero delay values at handoff to ZTL05 sector. To achieve these times, they modified the clearances recommended by the automation. A possible tool modification would be to allow the controllers to set goals for delay values along various routes and "tune" the automation accordingly. For example, perhaps the controller wants to hand off aircraft coming from the west with 20-30 seconds of delay, but would like to leave 40-50 seconds of delay for aircraft along the northwest flow. In the experiment, the controllers

often adjusted the delay value goal differently for the two main flows; it would be more useful for the automation to allow such offsets by route or flow (here, northwest flow and west flow).

First, by allowing the controller to input a set amount of expected delay (other than zero) to the trial planning tools, the controller could dynamically select the target delay based on observations of the actual errors, and thus flexibly customize the feed to the next sector based on the current conditions. The trial planning tools would resolve to this target delay time, so no manual adjustments would be needed.

Second, by allowing the controller to choose the target point, he/she could work to a point that aligns best with the responsibilities and airspace of the sector, instead of always working to the meter fix, which is often outside their sector (as it was for this experiment's ZTL06 controllers). For the high-altitude controller, the target point could be the point when the aircraft crosses into the next sector, since that is the level within their control (instead of to a zero delay at the meter fix). The meter fix should remain an often-chosen target point for the low-altitude controller. This is in line with the real-world interaction between en route sectors, and between en route and terminal airspace, where in each case, the controller attempts to provide a well-managed traffic flow ("feed") to the next controller. It allows flexibility for controllers to select the level of delay needed, in order to adjust to the level of error in the automation tools, such as a mismatch between forecast and actual winds. Assigning a custom point and measuring to it dynamically is very similar to how the Continuous Range Readout (CRR) function works in today's fielded en route system, where a distance countdown to a controller-specified point in space is shown by the aircraft position symbol.

Third, the controller could select different target delays for different flows. For example, in this experiment, such a capability would have allowed the controller to select a target delay value of -30 seconds for aircraft following the northwest flow, and -10 seconds for aircraft following the west flow. This is particularly helpful when the source of the trajectory uncertainty (perceived by the controllers as automation error) is due to a difference in winds, because based on wind direction, uncertainties will affect aircraft on different routes differently.

Controllers reported that the delay time updating immediately based on amendments to the system was very helpful, and they rated it as one of three tools they used all the time. It is interesting that the delay time in the datablock updating immediately was such a useful tool, because the trial planner shows the estimated delay if following the trial plan recommendations. Perhaps having the delay time persistently displayed served the controllers as a memory aid. Delay time in the datablock (that updates immediately based on trajectory change) is certainly a candidate for inclusion in future automation system enhancements; it was considered by the controllers to be much more useful in this experiment than in previous experiments simply because it was dynamically updated based on the trajectory change, instead of updating slowly based on aircraft movement.

By incorporating user input such as a route-specific, dynamically updatable target delay time and a user-configurable metering point along the arrival route, the automation would better support controllers in a flexible way, accounting for changing conditions and for varying controller strategies. If the work the controllers performed in manually adjusting the speeds recommended by the speed trial plan tool were allocated to the automation, the controllers could use the tool information directly, and they would be alleviated from the work of mentally adjusting the tool recommendations.

V. Conclusion

Controllers presented with the same traffic, tools, environment and goals independently selected very different control strategies for managing arrivals into Atlanta. Although they employed different strategies, both controllers were very successful at achieving the set goal – to absorb delay and combat the effects of both wind and aircraft performance errors to safely deliver aircraft to the meter fix on time.

Comparison of experimental data on the two controllers' strategies to working the same simulated arrival traffic in the same environment quantified how the two approaches differed and illustrated that the controllers had different ways of using the automation tools. Differences based on experiment condition were minimal but differences based on the amount of delay to absorb were shown. The two controllers' strategies entailed different types of clearances, which resulted in different lateral paths and vertical profiles when managing the arrival flows. The World 2 controller's strategy included issuing more clearances, and a wider variety of clearance types. The controllers were similarly effective in terms of schedule conformance, and flight path efficiency.

It is recommended that the FAA consider improvements to leverage the tools that both controllers used most often and reported favoring: the immediate update to the delay time in the datablock and having automation-provided speed suggestions. Additionally, allowing controllers to specify their own target delay times at controller-selected points dynamically based on environmental conditions is a way to allocate some of the manual work to the automation.

Acknowledgments

This work was conducted as an extension of the Functional Allocation and Separation Assurance work conducted in NASA Ames' Airspace Operations Laboratory. The experiment reported in this paper was sponsored by the NASA Airspace System Program, Concept and Technology Development Project. The authors wish to acknowledge Ashley Gomez, Connie Brasil, Faisal Omar, Vick Kelkar, and the AOL software team for their significant contributions to the experiment.

References

- ¹Federal Aviation Administration. "FAA Aerospace Forecast, Fiscal Years 2013-2033." FAA, Washington, D.C., 2013.
- ²A. Kirlik, "Modeling strategic behavior in human-automation interaction: Why an "aid" can (and should) go unused," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 35, p. 221242, 1993.
- ³S. M. Casner, "Understanding the determinants of problem-solving behavior in a complex environment," *Human factors: The Journal of the Human Factors and Ergonomics Society*, vol. 36, pp. 580-595, 1994.
- ⁴T. Prevot, "Exploring the Many Perspectives of Distributed Air Traffic Management: The Multi Aircraft Control System MACS," in *HCI-02 Proceedings*, Edmonton, 2002.
- ⁵J. Mercer, S. M. Green, T. Prevot, N. Beinert, A. Gomez, J. Kraut, S. Morey and G. Wu, "The Impact of Trajectory Prediction Uncertainty on Air Traffic Controller Performance and Acceptability," in *Aviation*, Los Angeles, 2013.
- ⁶J. Kraut, S. Morey, J. Mercer, A. Gomez and J. Homola, "How do Air Traffic Controllers Use Automation and Tools Differently During High Demand Situations?," in *Aviation*, Los Angeles, 2013.